



**DE LA SALLE UNIVERSITY**  
**College of Science**  
**Mathematics and Statistics Department**



**CATDATA – Categorical Data Analysis**  
*Prerequisite: LINEMOD*

*Prerequisite to:*

**Instructor:**  
**Consultation Hours:**

**Contact details:**  
**Class Schedule and Room:**

**Course Description**

A course on statistical methods for categorical (or cross-classified) response data.

**Learning Outcomes**

On completion of this course, the student is expected to present the following learning outcomes in line with the Expected Lasallian Graduate Attributes (ELGA) and the outcomes prescribed by the CHED Memorandum Order for the BS Mathematics program.

ELGA	Learning Outcome	Program Outcome													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Critical and Creative Thinker Effective Communicator Lifelong Learner	At the end of the course, the student will														
	perform tests of association between categorical variables	✓	✓	✓	✓	✓	✓			✓		✓	✓	✓	✓
	build models and assess their fit for categorical variables and draw conclusions from them	✓	✓	✓	✓	✓	✓			✓		✓	✓	✓	✓

**Program Outcomes (BS Statistics)**

A graduate of the program should be able to

1. Demonstrate broad and coherent knowledge and understanding of the core areas of statistical theory and statistical modeling .
2. Apply critical and problem solving skills using the scientific method.
3. Interpret scientific data and make judgments that include reflection on relevant scientific and ethical issues.
4. Carry out basic mathematical and statistical computations and use appropriate technologies in (a) the analysis of data; and ( b) In pattern recognition, generalization, abstraction, critical analysis and problem solving.
5. Communicate information, ideas problems and solutions, both, orally and in writing, to other scientists, decision makers and the public.
6. Relate science and mathematics with other disciplines.
7. Design and perform safe and responsible techniques and procedures in laboratory or field practices.
8. Critically evaluate input from others.
9. Appreciate the limitations and implications of science in everyday life.
10. Commit to the integrity of data.
11. Demonstrate broad and coherent knowledge and understanding in the core areas of statistics, computing and mathematics.
12. Generate information involving the conceptualization of a strategy for generating timely and accurate/reliable data, organizing a process for putting together or compiling the needed data, and transforming available data into relevant and useful forms.
13. Translate real-life problems into statistical problems.
14. Identify appropriate statistical tests and methods and their proper use for the given problems, select optimal solutions to problems and make decision in the face of uncertainty.

### Final Course Output

As evidence of attaining the above learning outcomes, the student is required to submit the following during the indicated dates of the term.

Learning Outcome	Required Output	Due Date
At the end of the course, the student will apply appropriate statistical concepts, processes, tools, and technologies in the solution to various conceptual and real-world problems.	An inquiry-based group presentation highlighting the uses of categorical data analysis in different problem situations encountered in the medical and related fields	Week 13

### Rubric for assessment

The following rubric will be used for grading students' rewritten solutions. The new quiz score will be obtained by adding ORIGINAL QUIZ SCORE and 20% of the REWRITTEN SCORE. Note that students will only rewrite items that they did not get perfectly.

CRITERIA	Excellent (4)	Good (3)	Satisfactory (2)	Needs Improvement (1)
<b>Formulation of the Research Problem and Objectives (10%)</b>	Research problem and objectives are clearly defined and significant; Demonstrates evidence that the research problem was researched and designed well.	Research problem and objectives are clearly defined and significant.	Research problem is clearly defined but some objectives are insignificant.	Research problem and objectives are vague and insignificant.
<b>Correct Application of the Statistical Concepts (35%)</b>	Statistical analyses are appropriate with correct interpretations and relevant conclusions.	Statistical analyses are appropriate with correct interpretations.	Some statistical analyses are inappropriate.	Statistical analyses are inappropriate
<b>Depth of Analysis (30%)</b>	The analysis convinces the reader about the wisdom of conclusions, implications and consequences on the basis of statistical methods and findings	The analysis engages the reader to appreciate the wisdom of conclusions, implications and consequences on the basis of statistical methods and findings	The analysis have limited ideas that do not explain the wisdom of conclusions, implications and consequences on the basis of statistical methods and findings	The analysis has incorrect ideas and conclusions.
<b>Clarity and Organization of Written Report (10%)</b>	Written report is organized logically and presented clearly with effective transitions.	Written report is organized logically and presented clearly.	Written report is organized and some discussions are not clear.	Written report is not organized.
<b>Oral Presentation (15%)</b>	Overall presentation is creative and well organized with innovative ideas.	Overall presentation is creative and well organized.	Overall presentation is organized	Overall presentation is not organized

### Additional Requirements

- Quizzes
- Class Participation (seatwork and group exercises, homework, recitation)
- Computer hands-on exercises using SAS
- Final Examination

### Grading System

	FOR EXEMPTED STUDENTS (w/out Final Exam)	FOR STUDENTS with FINAL EXAM	
		with no missed quizzes	with one missed quiz
Average of quizzes	90%	60%	50%
Other requirements	10%	10%	10%
Final exam/Project		30 %	40%

### Scale:

95-100%	4.0
89-94%	3.5
83-88%	3.0
78-82%	2.5
72-77%	2.0
66-71%	1.5
60-65%	1.0
<60%	0.0

### Learning Plan

LEARNING OUTCOME	TOPIC	WEEK NO.	LEARNING ACTIVITIES
At the end of the course, the student will apply appropriate statistical concepts, processes, tools, and technologies in the solution to various conceptual and real-world problems	1. Introduction 1.1 Levels of measurement 1.2 Categorical response data 1.3 Types of medical studies 1.4 Vital statistics and demographic methods	Week 1	Prior knowledge and beliefs survey Concept mapping Library work Group discussion and presentations
	2. Contingency Tables 2.1 Notation and Probability structure 2.2 Joint, Marginal, and Conditional probabilities 2.3 Independence 2.4 Comparing proportions 2.5 Relative risk and Odds ratio 2.6 Evaluation of a screening test	Week 2 - 3	Computer laboratory activities (SAS) Skills exercises Student self-assessment and reflection
	3. Inference for Two-Way Contingency Tables 3.1 CIs, SHyTs for proportions, RR, and OR 3.2 Chi-squared tests of independence 3.3 Fisher's exact test 3.4 Contingency coefficients	Week 4 - 5	
	<b>Quiz No. 1</b>	<b>Week 6</b>	
	4. Generalized Linear Models (GLMs) 4.1 Components of a GLM 4.2 Logit and Probit models 4.3 Poisson, Negative Binomial, and Count regression models	Week 6	
	5. Logistic Regression 5.1 Sampling models and their distributions 5.2 Fitting logit models 5.3 Odds ratio interpretation 5.3 Inference for logit models 5.4 Logit models with categorical predictors 5.5 Multiple logistic regression	Week 7 - 9	

	<b>Quiz No. 2</b>	<b>Week 9</b>	
	6. Loglinear Models for Contingency Tables 6.1 Three-way contingency tables 6.2 Types of independence: mutual, joint, conditional, marginal 6.3 Simpson's paradox 6.4 Abbreviated notation 6.5 Collapsibility 6.6 Model selection and diagnostics 6.7 Higher-dimensional contingency tables 6.8 The logistic-loglinear connection	Week 10 - 12	
	<b>Quiz No. 3</b>	<b>Week 12</b>	
	Group Reports* Final Project	Week 13	
	<b>Final Examination</b>	<b>Week 14</b>	

**\*Suggested topics for group reports:**

- Multicategory logit models for nominal responses
- Multicategory cumulative logit models for ordinal responses
- Paired-category ordinal logits
- Logistic regression for matched pairs

\*Skills exercises/ computer laboratory activities are given weekly and the students are expected to work on the solutions for their fourth hour activity. At the end of the term, the solutions to the problems will be compiled and submitted as one of the course outputs.

### References

- Agresti, Alan. (2007). *An Introduction to Categorical Data Analysis (2<sup>nd</sup> edition)*. Wiley.
- Agresti, Alan. (2013). *Categorical Data Analysis (3<sup>rd</sup> edition)*. Wiley.
- Fienberg, Stephen. (2004). *The Analysis of Cross-Classified Data, (2<sup>nd</sup> edition)*. Springer.
- Hosmer and Lemeshow (2000). *Applied Logistic Regression, (2<sup>nd</sup> Edition)*. Wiley-Interscience Publication.

### Online Resources

- GraphPad Software QuickCalcs from <http://www.graphpad.com/quickcalcs/CatMenu.cfm>
- Thompson, Laura A. (2008). R (and S-PLUS) Manual to Accompany Agresti's Categorical Data Analysis (2<sup>nd</sup> edition) Accessed September 26, 2012 from:  
<http://www.stat.purdue.edu/~zhanghao/MAS/handout/Splusediscrete2.pdf>
- Website for Agresti's Categorical Data Analysis, 2nd edition Accessed September 26, 2012 from:  
<http://www.stat.ufl.edu/~aa/cda2/cda.html>

### Class Policies

1. The required minimum number of quizzes for a 3-unit course is 3, and 4 for 4-unit course. No part of the final exam may be considered as one quiz.
2. Cancellation of the lowest quiz is not allowed even if the number of quizzes exceeds the required minimum number of quizzes.
3. As a general policy, no special or make-up tests for missed exams other than the final examination will be given. However, a faculty member may give special exams for
  - A. approved absences (where the student concerned officially represented the University at some function or activity).
  - B. absences due to serious illness which require hospitalization, death in the family and other reasons which the faculty member deems meritorious.
4. If a student missed two (2) examinations, then he/she will be required to take a make up for the second missed examination.
5. If the student has no valid reason for missing an exam (for example, the student was not prepared to take the exam) then the student receives 0% for the missed quiz.
6. Students who get at least 89% in every quiz are exempted from taking the final examination. Their final grade will be based on the average of their quizzes and other prefinal course requirements. The final grade of exempted students who opt to take the final examination will be based on the prescribed computation of final grades inclusive of a final examination. Students who missed and/or took any special/make-up quiz will not be eligible for exemption.
7. Learning outputs are required and not optional to pass the course.
8. Mobile phones and other forms of communication devices should be on silent mode or turned off during class.
9. Students are expected to be attentive and exhibit the behavior of a mature and responsible individual during class. They are also expected to come to class on time and prepared.
10. Sleeping, bringing in food and drinks, and wearing a cap and sunglasses in class are not allowed.
11. Students who wish to go to the washroom must politely ask permission and, if given such, they should be back in class within 5 minutes. Only one student at a time may be allowed to leave the classroom for this purpose.
12. Students who are absent from the class for more than 5 meetings will get a final grade of 0.0 in the course.
13. Only students who are officially enrolled in the course are allowed to attend the class meetings.

Approved by:



**DR. JOSE TRISTAN F. REYES**  
Chair, Mathematics and Statistics